PIBB 生物化学与生物物理进展
Progress in Biochemistry and Biophysics
2015, 42(4): 356~364
www.pibb.ac.cn

凝集素法与酰肼法富集糖肽的方法学比较*

薛　鹏[1,2)]　周　岳[1,2,3)]　丁　翔[1)]　王　峻[1,2)]　谢振声[1)]　杨福全[1)**]

([1)] 中国科学院生物物理研究所，蛋白质组学技术实验室，北京 100101；[2)] 中国科学院大学，北京 100049；
[3)] 赛默飞世尔科技有限公司，上海 200120)

摘要　蛋白质糖基化(glycosylation)是最常见和最重要的翻译后修饰之一．大规模 N- 连接糖基化位点鉴定是糖蛋白质组学研究的重要组成部分，而 N- 连接糖肽富集是高通量 N- 连接糖基化位点鉴定的关键步骤．凝集素富集法和酰肼化学法是目前被广泛应用的 N- 连接糖肽富集技术，有报道认为两种方法具有很强的互补性，联合使用能提高糖基化位点的鉴定数目．本文以 HepG2 细胞系为模型，系统比较了这两种方法的富集效率和糖蛋白鉴定数目．结果表明，虽然酰肼法的糖肽富集效率为 76.6%，远高于凝集素法的 54.6%，但是凝集素法却能鉴定到 825 个糖蛋白和 1 959 个 N- 连接糖基化位点，显著多于酰肼法富集到的 522 个糖蛋白和 1 014 个糖基化位点．并且，两种方法并未显示出显著的互补性，仅 28 个糖蛋白和 80 个糖基化位点未在凝集素法中鉴定到．

关键词　N- 连接糖基化蛋白质组学，凝集素富集法，酰肼化学富集法
学科分类号　O658，Q51　　**DOI**: 10.16476/j.pibb.2014.0324

蛋白质糖基化对于蛋白质的折叠、运输和定位等都起着重要作用，并参与受体激活、信号转导、细胞 - 细胞相互作用等诸多重要的生物学过程[1-4]．高通量鉴定 N- 连接糖蛋白以及糖基化位点是糖蛋白质组学研究的重要组成部分．尽管超过 50%的哺乳动物蛋白质发生了糖基化，然而糖基化肽段仅占到蛋白质酶解肽段的 2%～5%[5]，对于复杂的生物样本，蛋白质含量动态范围宽，大量高丰度非糖基化修饰蛋白的存在，使得低丰度糖基化修饰蛋白检测更加困难，另一方面，就同一蛋白来说，糖基化位点和修饰有着高度的不均一性，增加了糖基化肽段检测的复杂度，因此，必须对糖蛋白或糖肽进行富集以提高检测灵敏度．

在 N- 连接糖蛋白质组学研究中，通常先将蛋白质样品酶解成肽段后进行糖肽富集，再用 PNGase F 酶切除糖肽上的糖链，此时糖基化修饰的天冬酰胺残基转化为天冬氨酸并产生+0.984Da 的质量偏移．经液相色谱 - 串联质谱分析后，以天冬酰胺转化为天冬氨酸作为可变修饰进行数据库检索，便能得到糖蛋白及其糖基化位点信息[6]．目前已报道的糖肽富集方法有酰肼化学富集法[7-8]、凝集素富集法[9-11]、二氧化肽(TiO_2)亲和富集[12]、硼酸富集[13]、亲水相互作用色谱富集法[14]等．其中最为常用的是酰肼法和凝集素法．酰肼法是利用糖蛋白或糖肽糖链上的顺式邻二羟基结构(cis-diols)与高碘酸反应生成醛，醛基在与酰肼树脂上的氨基反应形成共价的腙键，从而被酰肼树脂捕获，PNGase F 酶将糖肽从酰肼树脂上释放，得到富集后的糖基化肽段．2012 年，Sudhir[8]等运用酰肼化学富集法在人支气管上皮细胞(HBECs)中鉴定到了 317 个 N- 连接糖蛋白含 608 个 N- 连接糖基化位点．凝集素法是通过凝集素选择性地识别某些特殊结构的单糖或聚糖实现对糖蛋白的富集．不同糖型的糖肽可以采用不同的凝集素富集，刀豆凝集素(convalina agglutinin，Con A)富集高甘露糖型的糖肽、麦胚凝集素(wheat germ agglutinin，WGA) 富集末端含有 N- 乙酰氨基葡萄糖和唾液酸的糖肽、蓖麻凝集素(ricinus communis agglutinin Ⅰ，RCA120)能富集末

* 国家重点基础研究发展计划(973)(2012CB966803, 2010CB833703)和中国科学院仪器设备功能开发技术创新项目(yg201001)资助.
** 通讯联系人.
Tel: 010-64888581，E-mail: fqyang@ibp.ac.cn
收稿日期：2015-01-20，接受日期：2015-03-17

端含有β1，4-连接半乳糖的糖肽、橙黄网胞盘菌凝集素(aleuria aurantia lectin，AAL)能富集岩藻糖的糖肽、无梗接骨木凝集素(sambucus sieboldiana agglutinin，SSA)、黑接骨木凝集素(sambucus nigra agglutinin，SNA)、怀槐凝集素(maackia amurensis lectin，MAL)都能富集唾液酸的糖肽．Mann研究团队[11]联合使用3种凝集素(ConA、WGA和RCA120)，利用超滤辅助的样本制备方法(filter aided sample preparation，FASP)，对小鼠的血浆、脑、肾脏、肝脏、心脏组织的N-连接糖基化位点进行了全面的解读，共鉴定到来自2 352个糖蛋白的6 367个N-糖基化位点．Kaji等[15]使用5种凝集素(ConA、RCA120、WGA、AAL和SSA)并优化了不同凝集素的组合，同时结合亲水相互作用色谱法(Amide-80)，同样对小鼠的多种组织进行检测，富集到2 556个糖蛋白的5 657个糖基化位点．酰肼化学法和凝集素法对N-连接的糖蛋白的富集原理不同，一些研究将两种富集方法联合使用，并显示出一定的互补性．McDonald等[16]只采用凝集素MAL在HeLa细胞膜上富集到45个糖基化蛋白，酰肼法鉴定到56个糖基化蛋白，两种方法共鉴定到73个糖蛋白．Song等[17]联合两种方法在食道癌病人的血浆中共鉴定到139个糖蛋白，其中两种方法共同鉴定到的糖蛋白数目仅为59个．本文用酰肼法和多种凝集素富集法对HepG2细胞系的N-连接糖肽进行了富集，利用LTQ Oribtrap XL[18]串联质谱鉴定到825个N-连接糖蛋白相对应1959糖基化位点，结果显示凝集素富集法基本上覆盖了酰肼化学富集法鉴定到的N-连接糖蛋白和糖基化位点．

1 实验部分

1.1 实验试剂和仪器

高碘酸钠($NaIO_4$)、酰肼树脂(hydrazide resin)购自Bio-Rad公司(Hereules，CA，USA)．碘乙酰胺(iodoacetamide，IAA)、二硫苏糖醇(dithiothreitol，DTT)、甲酸、尿素、硫脲和离心柱购自Sigma公司(Milwaukee，WI，USA)．固相萃取柱Oasis HLB购自美国Waters公司(Milford，USA)．糖肽内切酶PNGase F(500U/μL)购自New England Biolabs公司(Ipswieh，MA，USA)．测序级胰蛋白酶购自Promega公司(Madison，WI，USA)．测序级蛋白酶Lys C购自Wako公司(Osaka，Japan)．Con A、WGA和RCA 120购自Vector labs公司(Burlingame，CA，USA)．

PolySULFOETHYL A强阳离子交换(SCX)固相萃取柱购自PolyLC公司(Columbia，MA，USA)．实验室自制的捕集柱和分析柱填料(C18 Kinetex 2.6 μm)购自于Phenomenex公司(Torrance，CA，USA)．纳升液相色谱Nano LC 1D plus购自Eksigent公司(Redwood，CA，USA)．线性离子阱静电场轨道阱质谱仪(LTQ-Orbitrap XL)购自Thermo公司(Bremen，Germany)．

1.2 样品制备

HepG2细胞采用青霉素链霉素双抗的DMEM培养基，37℃，5% CO_2恒温培养箱培养．待生长到超过培养皿90%时，用1 ml 10 mmol/L PBS (pH 7.5)清洗3次，然后加入1 ml裂解液(含7 mol/L尿素和2 mol/L硫脲)，冰上微摇裂解，超声破碎，离心去除沉淀，用Bradford法测定蛋白质浓度，储存于−80℃冰箱备用．取12 mg蛋白样品加入终浓度10 mmol/L二硫苏糖醇(DTT)，室温下振荡反应45 min，进行蛋白质二硫键还原；然后加入30 mmol/L碘乙酰胺(IAA)，于室温和避光条件下孵育45 min，使巯基烷基化；待反应结束后向溶液加入DTT至终浓度5 mmol/L以除去过量的IAA；按蛋白酶与蛋白1∶200的质量比加入Lys C，37℃酶解2 h，用50 mmol/L NH_4HCO_3缓冲液溶解稀释4倍，按蛋白酶与蛋白1∶100的质量比加入Trypsin，37℃酶解过夜．酶解液经Oasis HLB柱脱盐．

1.3 强阳离子交换(SCX)小柱预分离

将上述脱盐后的多肽混合样品用SCX小柱预分成9个组分．流程如下：先用7 mmol/L KH_2PO_4、30%乙腈(ACN)缓冲液(pH 2.65)润洗SCX固相萃取柱，然后将12 mg多肽混合样品反复3次上样，再分别用含有10、20、30、40、60、90和350 mmol/L KCl及7 mmol/L KH_2PO_4、30% ACN的缓冲液(pH 2.65)依次洗脱并收集相应的组分．预分后的各个组分浓缩除去乙腈后，再用Oaiss HLB柱脱盐．

1.4 凝集素富集

将预分各组分肽段分别溶解在500 μl 50 mmol/L NH_4HCO_3溶液中，然后加入ConA、WGA和RCA120三种凝集素偶联的琼脂糖颗粒，在4℃下孵育1 h；用500 μl含有50 mmol/L NH_4HCO_3的10% ACN溶液洗凝集素颗粒3次，除去非特异性吸附肽段；再用200 μl 50 mmol/L NH_4HCO_3重悬

凝集素颗粒，100℃加热 5 min 后收集上清，并加入 1 μl PNGase F 酶(500U/μl)，在 37℃下孵育 4 h 后离心收集脱糖基化的肽段.

1.5 酰肼法富集

将预分的各组分肽段分别溶解于 500 μl 偶联缓冲液中 (100 mmol/L乙酸钠，150 mmol/L氯化钠，pH 5.5)；加入新鲜配制的高碘酸钠，使其终浓度为 10 mmol/L，反应 1 h；加入新鲜配制的亚硫酸钠，使其终浓度为 20 mmol/L，振荡反应 10 min，消耗过量的氧化剂. 向肽段溶液中加入已用偶联缓冲液润洗 5 遍的酰肼树脂，在 37℃条件下进行偶联反应，振荡 12 h 使树脂与溶液充分混合；依次用 1 ml 1.5 mol/L氯化钠溶液、100%甲醇、乙腈和 50 mmol/L NH_4HCO_3 各清洗 3 次以洗去树脂上的非特异吸附肽段；加入 100 μl 50 mmol/L NH_4HCO_3 溶液和 1 μl PNGase F(500U/μl)，在 37℃下孵育 4 h 后离心收集脱糖基化的肽段.

1.6 糖基化位点假阳性率的测定

为了评估酶解、富集和脱糖过程中化学脱氨对糖基化位点鉴定的影响，我们分别对凝集素法和酰肼法进行了假阳性率分析. 采用相同的 HepG2 酶解样品，与前面描述的富集方法类似，只是凝集素富集时糖肽不采用 PNGase F 处理，酰肼法富集前多肽不经过高碘酸氧化处理[19]. 这样得到的多肽不具有糖基化位点，而具有 NxS/T(x≠P)序列特征的多肽属于化学脱氨肽段. 这些具有 NxS/T(x≠P)糖基化序列特征的化学脱氨肽段与具有 NxS/T(x≠P)序列特征正常富集到的糖基化肽段的百分比即为糖基化位点的假阳性率.

1.7 NanoHPLC-MS/MS 分析

将富集后的肽段以 2 μl /min 流速上到捕集柱(20 mm×150 μm)，再用 250 nl/min 流速洗脱到分析柱(15 cm×75 μm)分析. 色谱条件为：流动相 A，0.5% 甲酸(FA)水溶液；流动相 B，0.5% FA 的乙腈溶液. 色谱梯度为：0 min(0%B)-10 min(12%B)-140 min (33% B)-155 min (100% B)-165 min (100% B). 串联质谱采用线性离子阱静电场轨道阱质谱仪(LTQ-Orbitrap XL)，一级质谱采用 Orbitrap 检测，分辨率 m/z 400 处为 60 000；采集质量范围 m/z 300～2000，依次选取一级质谱中离子强度最强的 10 个离子进行 CID 二级串联质谱分析.

1.8 数据库检索及功能分析

质谱原始数据用 Maxquant 1.3.05[20]进行检索，数据库为 UniProt 人类蛋白质数据库(版本 2012 年 11 月，89601 条蛋白序列). 数据库搜索参数为：一级谱中母离子质量误差为±20，二级谱中碎片离子质量误差为±0.8Da，酶为胰蛋白酶(trypsin)，完全酶切，漏切数目为 2 个，固定修饰为半胱氨酸的烷基化(C，+57.02150Da)，可变修饰为天冬氨酰胺脱酰胺基(N，+0.98402Da)和甲硫氨酸的氧化(M，+15.99492Da). PSM (peptide spectrum match)、肽段、蛋白质和天冬氨酰胺脱酰胺基位点的假阳性率均控制在 1%，鉴定肽段的最小长度为 6 个氨基酸. 当天冬酰胺的脱酰胺基肽段符合 NxS/T(x≠P)序列特征，被确认为 N- 连接糖基化肽段. 糖蛋白的 GO 和 KEGG 功能分析采用 Persus1.3.0.4 软件.

2 结果与讨论

2.1 实验设计策略

本研究考察凝集素富集法和酰肼化学法对 N-连接糖肽的富集差异时，首先对 HepG2 细胞系多肽酶解产物用强阳离子交换小柱预分，再分别对各个组分富集，从而大规模的获得糖肽. 如图 1 所示，肽段在预分前和预分后进行了 HLB 脱盐，既减少盐离子对糖肽富集的影响，又有助于强阳离子交换柱的分离. 然而，糖基的亲水性导致部分糖肽的亲水性增强，在 HLB 保留减弱. 因此，本研究也将预分前 HLB 脱盐的流穿液(flow through)作为

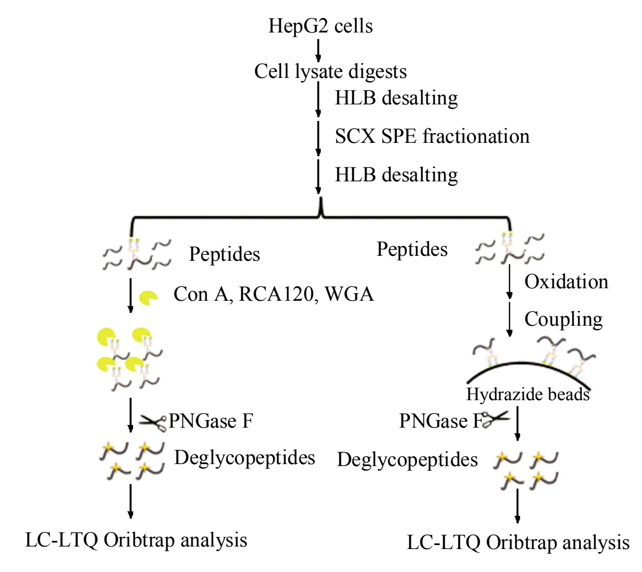


Fig. 1 A schematic overview of lectins (left) and hydrazide chemistry (right) based enrichment methods for N-linked glycoproteomics

The lysates of HepG2 cells were tryptic digested, then desalted with Oasis HLB column and separated into 9 fractions by SPE-SCX. Each fraction was quartered and enriched by hydrazide chemistry and lectins respectively. Enriched N-glycopeptides were digested with PNGase F to remove glycans and then analyzed by LTQ Orbitrap XL.

一个组分 HLB-FT，连同 SCX 流穿液(flow through)组分 SCX-FT， SCX 预分离的 7 个组分 10、20、30、40、60、90和 350 mmol/L KCl，分别进行了凝集素法和酰肼法两次重复富集．凝集素法按一定比例联合使用 ConA、WGA 和 RCA120 三种凝集素，能够特异性地富集高甘露、N- 乙酰氨基葡萄糖和唾液酸、β1, 4- 连接半乳糖等不同糖型的糖肽，提高糖蛋白的鉴定种类．酰肼化学法富集时采用经典的富集步骤．两种方法富集得到的糖肽均用 PNGase F 切除糖链，采用 LTQ Orbitrap XL 高精度质谱 150 min 有效梯度检测．采用 Maxquant 软件分析，并筛选含有 NxS/T 序列特征的位点作为 N-连接糖基化位点．

2.2　凝集素法和酰肼法鉴定糖基化位点数目的比较

凝集素法和酰肼法总共富集到 853 个 N- 连接糖蛋白和 1 959 个糖基化位点．如图 2a 所示，采用三种凝集素联合富集的方法鉴定到了 3 047 个天冬酰胺脱酰胺基修饰肽段合计 2 507 个脱酰胺基位点，其中具有 NxS/T(x≠P)序列特征的糖基化位点 1 879 个，匹配到 825 个糖蛋白．而酰肼化学法仅鉴定到 1 346 个天冬酰胺脱酰胺基修饰肽段合计 1 320 个脱酰胺基位点，其中 1 014 个为糖基化位点，对应 522 个糖蛋白．无论是天冬酰胺脱酰胺肽段和脱酰胺位点，还是糖基化蛋白和糖基化位点，凝集素富集法鉴定的数目均高于酰肼法．在所鉴定的天冬酰胺脱酰胺位点中，大约 25%的位点不符合 NxS/T(x≠P)糖基化序列特征．是由于在 Trypsin 和 PNGase F 的碱性酶解反应条件下，天冬酰胺(N)会自发脱酰胺[21]，这些脱酰胺位点在凝集素法或酰肼法的富集过程中作为非特异性吸附而被质谱检测到．为了检测自发脱氨对糖基化位点鉴定的影响，我们采用 Palmisano 等[22]提出的计算方法对凝集素法和酰肼法进行了假阳性率分析，其中凝集素法未经过 PNGase F 酶处理，鉴定到 114 个脱氨位点，23 个位点具有 NxS/T(x≠P)糖基化序列特征，假阳性率为 1.22%(23/1 879)．酰肼法未经 PNGase F 酶处理，得到了 55 个脱氨位点含有 8 个具有 NxS/T (x≠P)特征序列的位点，假阳性率为 0.79%(8/1 014).

本研究中，凝集素法富集的糖基化位点数目 1 879 个不仅比酰肼法富集糖基化数目 1 014 个多了近一倍，而且覆盖了 95.9%的总糖基化位点数，仅有 80 个糖基化位点是酰肼法特异鉴定到的(图 2b). 在糖蛋白水平具有相同的结论，凝集素鉴定到的蛋白数目涵盖了超过 96.7%总糖蛋白，仅有 28 个糖蛋白是在凝集素法未鉴定到的(图 2c)．而 McDonalde 曾报道凝集素富集法和酰肼化学富集法具有一定的互补性[15]，在 HeLa 细胞中共鉴定到 127 个 N- 连接糖基化位点，其中凝集素法富集鉴定到 76 个位点，酰肼法富集鉴定了 75 个．导致不同结论的原因可能与 McDonalde 使用单凝集素(MAL)富集策略有关．本研究采用了 ConA、WGA 和 RCA120 三种凝集素联合富集，同时辅以 SCX 预分，因此最大限度地提高了凝集素法富集鉴定糖肽的数目．

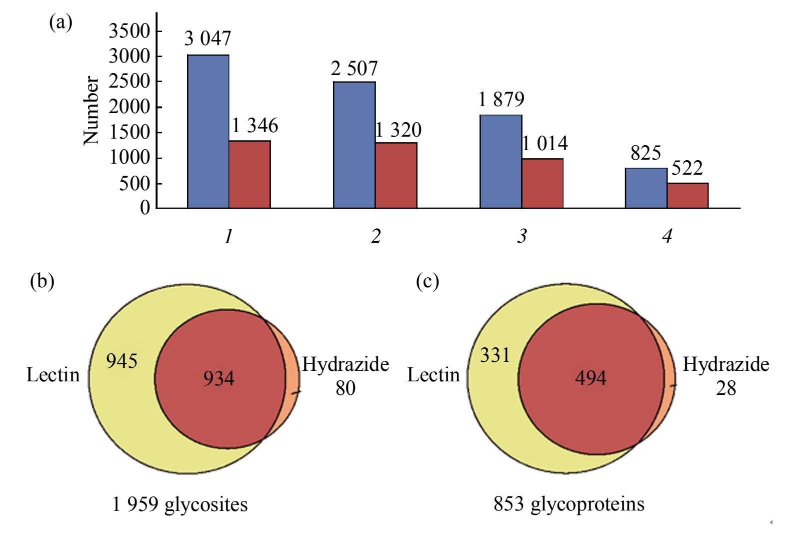


Fig. 2　Summary of the results from lectins and hydrazide based enrichment methods

(a) The number of deamidated peptides, deamidated sites, deamidated NxS/T (x≠P) sites and glycoproteins in lectins (blue) and hydrazide chemistry (red) based enrichment methods. *1*: Deamidated peptides; *2*: Deamidated sites; *3*: Deamidated sites with NxS/T motif; *4*: Glycoproteins. ■: Lectins; ■: Hydrazide. (b) Overlap of glycosites between lectins and hydrazide chemistry enrichment methods. (c) Overlap of glycoproteins between lectins and hydrazide chemistry enrichment methods.

图 3 显示，在所鉴定的 3 051 个 N- 连接糖基化肽段中，只有 1 个糖基化位点的糖肽超过 80%，2 个糖基化位点的糖肽约占 14%，3 个及以上糖基化位点的糖肽不足 2.5%(图 3a)．而在整个蛋白质水平，含有 1 个、2 个、3 个和 4 个糖基化位点的蛋白质则分别占 44.3%、20.0%、13.4%和 6.5% (图 3b)，与 Zielinska 等[11]报道的结果基本一致．

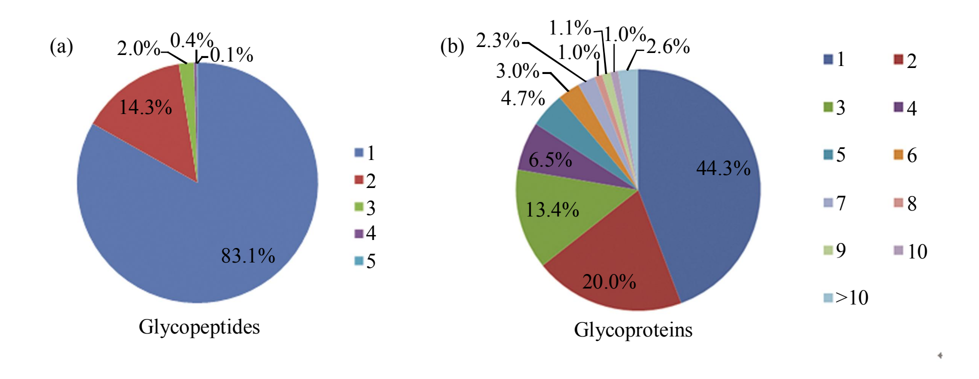


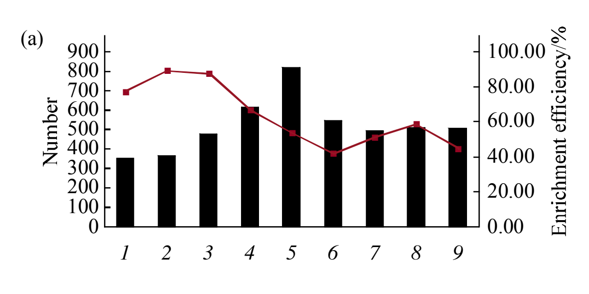
Fig. 3　Distribution of single and multiple glycosylated peptides and identified by lectin and hydrazide chemistry enrichment methods proteins

(a) Distribution of single and multiple glycopeptides. (b) Distribution of single and multiple glycosylated proteins.

2.3　凝集素法和酰肼法糖肽富集效率比较

糖基化肽段的富集效率是指鉴定到的糖基化修饰肽段占所有检测肽段的百分比．如图 4a 所示，HLB-FT、SCX-FT 及 10、20、30、40、60、90 和 350 mmo/L KCl 等 9 个组分，通过凝集素法富集鉴定的糖基化位点分别为: 355、367、481、617、818、649、497、615 和 507．富集效率分别为: 77.2%、89.3%、87.6%、67.1%、53.9%、41.9%、51.4%、58.9%和 44.9%．平均富集效率为 54.6%，略高于 Zielinska 等[11]报道的 45%．值得注意的是通常被忽略或丢弃的 HLB 流穿液(HLB-FT)、SCX 流穿液(SCX-FT)，以及低盐浓度洗脱(10 mmol/L KCl)三个组分的糖肽富集效率显著高于其他组分．这可能是由于糖基化肽段比其他非糖基化肽段的亲水性较强，在 HLB 反相捕集柱上保留较弱或无保留，使得强亲水性糖肽富集于 HLB 流穿液中．而含有唾液酸的酸性糖肽在 SCX 柱上不保留或保留较弱，因此在 SCX-FT 和 10 mmol/L KCl 两个组分中也有较高的含量， Lewandrowski 等[23]利用该原理采用 SCX 直接富集到 79 个含唾液酸的糖蛋白．

如图 4b，酰肼化学法的 9 个组分富集效率分别为 59.4%、81.3%、87.9%、80%、56.2%、70.5%、76.4%、70%和 78.6%，平均的富集效率高达 76.6%．酰肼法富集的糖基化位点数比凝集素法低，而富集效率却明显高于凝集素法．这可能是酰肼法的缩合反应要比凝集素糖的亲和反应特异性更高，并且本研究中酰肼法采用聚合物树脂为比凝集素的琼脂糖凝胶树脂非特异吸附更低，所以糖基化肽段富集效率高，但酰肼法富集过程中使用高碘酸，会造成肽段的酪氨酸残基碘化及其他一些化学



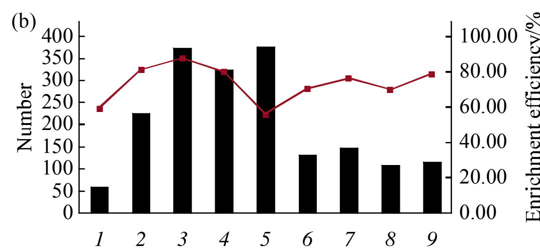


Fig. 4　Distribution and comparison of N-glycositess and enrichment efficiency in nine fractions

(a) Lectin method. (b) Hydrazide chemistry method. *1*: HLB-FT; *2*: SCX-FT; *3*: 10 mmol/L KCl; *4*: 20 mmol/L KCl; *5*: 30 mmol/L KCl; *6*: 40 mmol/L KCl; *7*: 60 mmol/L KCl; *8*: 90 mmol/L KCl; *9*: 350 mmol/L KCl. ■: Glycosites; ■—■: Enrichment efficiency.

副反应[24]，肽段因修饰而检索率降低，因此糖基化肽段的鉴定数目较少.

2.4 N-连接糖基化位点序列特征分析

糖基化位点序列特征统计分析显示，1 135个糖基化位点含有 NxT 特征序列，占位点总数的 57.9%，824 个糖基化位点含有 NxS 的特征序列，占位点总数的 42.1%. 这与 Wuhrer 报道相一致[11]. 应用 WebLogo[25]工具对氨基酸频次分析，结果也表明 NxT 序列特征出现频率显著高于 NxS 的出现频率(图 5a). 糖基化位点 N 后的第一个氨基酸 x(横坐标为 1)，在 NxT 特征序列中出现的最高频率是缬氨酸 V(图 5b)，而在 NxS 的特征序列中 x 为丝氨酸 S(图 5c). 非糖基化的脱酰胺基肽段(图 5d)与糖肽中氨基酸频次不同，这些自发脱酰氨天冬酰胺最常出现在甘氨酸 G、天冬氨酸 N 和丝氨酸 S 之前. 这与 Hao 等[21]报道的 -N-G-，-N-S- 和 -N-D- 等序列特征更容易在碱性条件下发生自发脱氨相一致.

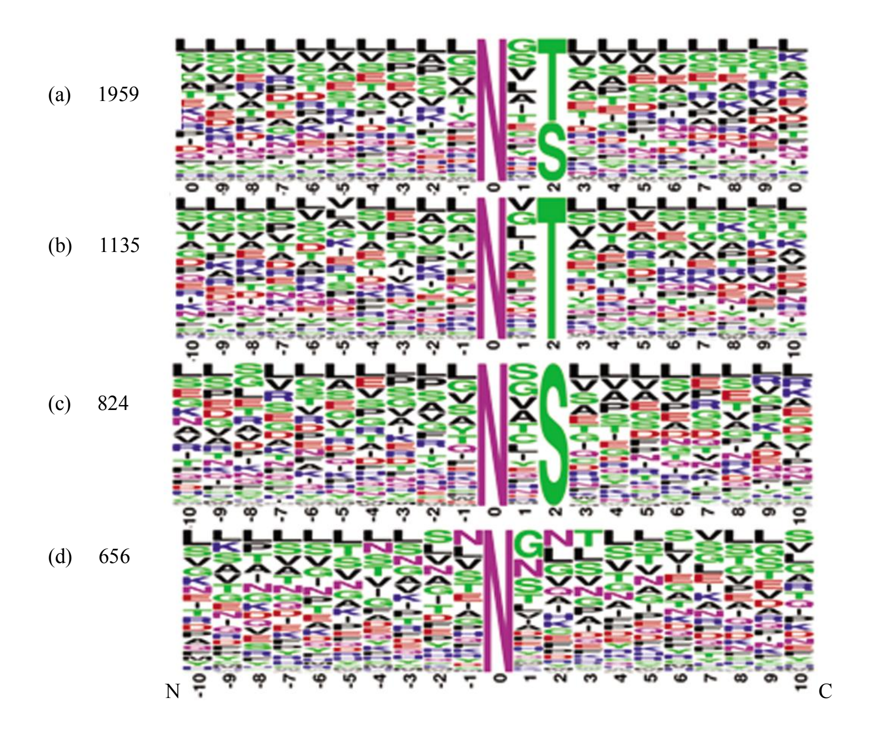


Fig. 5 Analysis of sequnence motif using WebLogo

(a) N-glycosylated consensus sequence with N-x-S/T (X ≠ P) motif, (b) N-glycosylated consensus sequence with N-x-T motif (c) N-glycosylated consensus sequence with N-x-S motif (d) Non-glycosylaited deamidated consensus sequence.

2.5 糖蛋白的生物信息分析

本研究在 HepG2 细胞中鉴定到 853 个糖蛋白，其中 593 个在 UniProt 数据库中注释为糖蛋白，260 个新发现的 N- 连接糖蛋白. 为了进一步了解糖蛋白的总体功能，我们进行了 GO (gene ontology)slim 的功能分析. 在细胞组成(cellular components)方面，糖蛋白主要集中分布在细胞外基质(17.24%)、细胞质膜(10.67%)、内质网 / 高尔基体(6.92%)等细胞膜相关体系(图 6a). 而在生物过程(biological processes)方面，糖蛋白也主要参与了信号转导(14.89%)、转运(13.72%)、蛋白翻译后修饰(9.96%)、细胞黏附(9.96%)、免疫应答(8.32%)等生物过程(图 6b). 同时 KEGG 通路分析表明，糖蛋白质主要参与了溶酶体、蛋白质内质网加工、肿瘤的信号转导、细胞因子相互作用、吞噬体、焦点黏附、ECM- 信号转导和细胞黏附信号转导通路(图 6c). 糖蛋白与细胞表面膜蛋白高度相关，而细胞表面分化抗原簇(cluster of differentiation antigens，CDs)是一类非常重要的细胞表面膜相关蛋白质家族[15, 26]. 如表 1 所示，我们共计发现了 72 个人类 CD 蛋白检测到 N- 连接的糖基化位点. 这暗示了糖基化修饰对 CD 的分子功能、识别和免疫信号的传导具有重要的作用.

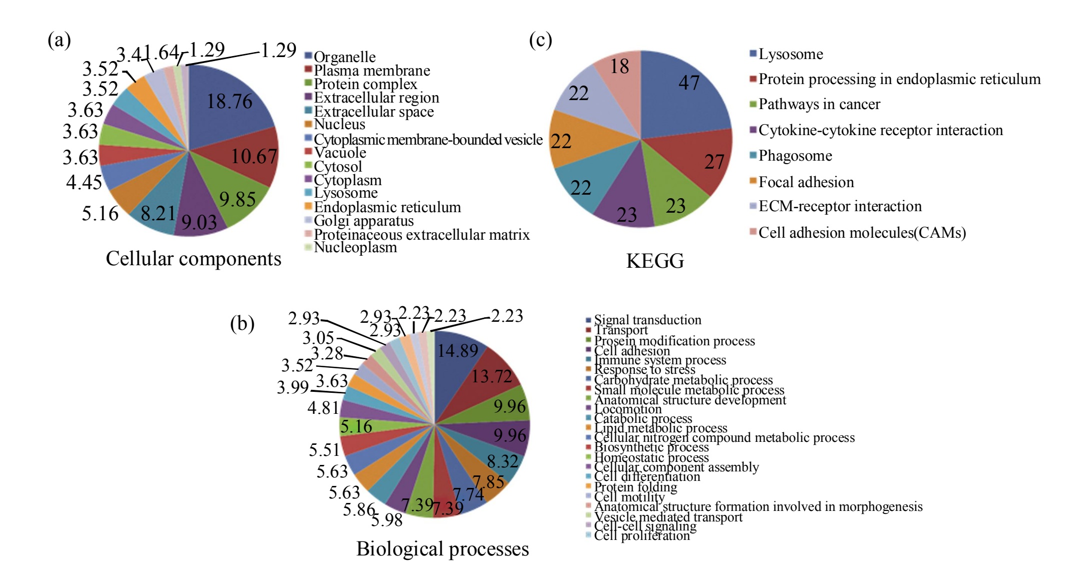


Fig. 6 Distribution of glycoproteins based on functional categories

(a) Cellular component. (b) Biological process. (c) KEGG. The number in (a) and (b) indicates the percentage of total glycoproteins belongs to a given functional category. The number in (c) indicates the number of identified glycoproteins in the given pathway.

Table 1 CD molecules from all identified glycoproteins

Gene ID	Uniport ID	CD name	Gene ID	Uniport ID	CD name
ABCB1	P08183	CD243	IL17RA	Q96F46	CD217
ABCG2	Q9UNQ0	CD288	IL6R	P08887	CD126
ADAM10	O14672	CD156c	IL6ST	P40189	CD130
ADAM17	P78536	CD156B	INSR	P06213	CD220
ALCAM	Q13740	CD166	ITGA1	P56199	CD49a
ATP1B3	P54709	CD298	ITGA2	P17301	CD49b
BCAM	P50895	CD239	ITGA5	P08648	CD49e
BST2	Q10589	CD317	ITGAV	P06756	CD51
CD109	Q6YHK3	CD109	ITGB4	P16144	CD104
CD151	P48509	CD151	L1CAM	F5H025	CD171
CD164	Q04900	CD164	LAMP1	P11279	CD107a
CD276	Q5ZPR3	CD276	LAMP3	Q9UQV4	CD208
CD33	P20138	CD33	LIFR	P42702	CD118
CD36	P16671	CD36	LRP1	Q07954	CD85l
CD44	P16070	CD44	MCAM	P43121	CD146
CD47	Q08722	CD47	MFI2	P08582	CD228
CD55	B1AP13	CD55	MRC2	Q9UBG0	CD280
CD58	P19256	CD58	MST1R	Q04912	CD136
CD59	E9PR17	CD59	NRP1	O14786	CD304
CD63	P08962	CD63	NT5E	P21589	CD73
CD68	P34810	CD68	PLAUR	Q03405	CD87
CD70	P32970	CD70	PRNP	P04156	CD230
CD82	P27701	CD82	PROCR	Q9UNN8	CD201
CD97	P48960	CD97	PTGFRN	Q9P2B2	CD315
CDCP1	Q9H5V8	CD318	PTPRJ	Q12913	CD148
CDH2	P19022	CD325	PVR	P15151	CD155
CDH5	P33151	CD144	PVRL1	Q15223	CD111
CSF1R	P07333	CD115	PVRL2	Q92692	CD112
TNFRSF1A	P19438	CD120a	PVRL3	Q9NQS3	CD113
EMR2	Q9UHX3	CD312	SEMA4D	Q92854	CD100
ERBB2	P04626	CD340	SEMA7A	O75326	CD108
F3	P13726	CD142	IL13RA1	P78552	CD213a1
FGFR4	P22455	CD307d	TFRC	P02786	CD71
GGT1	P19440	CD224	TLR6	Q9Y2C9	CD286
IGF1R	P08069	CD221	TNFRSF11A	Q9Y6Q6	CD265
IGF2R	P11717	CD222			
IGSF8	Q969P0	CD316			

3 结 论

本文以 HepG2 细胞为模型，比较了酰肼化学法和三种凝集素混合富集法鉴定 N- 连接糖基化位点数目和富集效率. 其中凝集素富集鉴定到 1 879 个 N- 连接糖基化位点、825 个糖蛋白质，酰肼化学法富集鉴定到 1 014 个 N- 连接糖基化位点、522 个糖蛋白质. 两种富集方法总共鉴定到 1 936 个 N- 连接糖基化位点，853 个糖蛋白. 虽然酰肼法在糖肽富集时表现出更好的特异性，但富集鉴定糖基化位点的数目显著低于凝集素法. 而且三种凝集素混合富集鉴定的糖基化位点和糖蛋白几乎涵盖了酰肼法的富集鉴定结果. 此外，在利用 HLB 柱脱盐和 SCX 柱预分时，其流穿液富含糖肽，应加以利用.

参 考 文 献

[1] Haltiwanger R S, Lowe J B. Role of glycosylation in development. Annu Rev Biochem, 2004, **73**: 491−537

[2] Hart G W, Copeland R J. Glycomics hits the big time. Cell, 2010, **143**(5): 672−676

[3] Ohtsubo K, Marth J D. Glycosylation in cellular mechanisms of health and disease. Cell, 2006, **126**(5): 855−867

[4] Zhang Y, Zhao J H, Zhang X Y, *et al*. Relations of the type and branch of surface N-glycans to cell adhesion, migration and integrin expressions. Molecular and Cellular Biochemistry, 2004, **260**(1−2): 137−146

[5] Sun B, Ranish J A, Utleg A G, *et al*. Shotgun glycopeptide capture approach coupled with mass Spectrometry for comprehensive glycoproteomics. Mol Cell Proteomics, 2007, **6**(1): 141−149

[6] Pan S, Chen R, Aebersold R, *et al*. Mass spectrometry based glycoproteomics——from a proteomics perspective. Mol Cell Proteomics, 2011, **10**(1): R110 003251

[7] Zhang H, Li X J, Martin D B, *et al*. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnology, 2003, **21** (6): 660−666

[8] Sudhir P R, Chen C H, Pavana Kumari M, *et al*. Label-free quantitative proteomics and N-glycoproteomics analysis of KRAS-activated human bronchial epithelial cells. Mol Cell Proteomics, 2012, **11**(10): 901−905

[9] Kaji H, Yamauchi Y, Takahashi N, *et al*. Mass spectrometric identification of N-linked glycopeptides using lectin-mediated affinity capture and glycosylation site-specific stable isotope tagging. Nat Protocol, 2006, **1**(6): 3019−3027

[10] Rüdiger H, Gabius H J. Plant lectins: occurrence, biochemistry, functions and applications. Glycoconj J, 2001, **18**(8): 589−613

[11] Zielinska D F, Gnad F, Wiśniewski J R, *et al*. Precision mapping of an *in vivo* N-Glycoproteome reveals rigid topological and sequence constraints. Cell, 2010, **141**(5): 897−907

[12] Larsen M R, Jensen S S, Jakobsen L A, *et al*. Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. Mol Cell Proteomics, 2007, **6**(10): 1778−1787

[13] Sparbier K, Wenzel T, Kostrzewa M. Exploring the binding profiles of ConA, boronic acid and WGA by MALDI-TOF/TOF MS and magnetic particles. J Chromatogr B, 2006, **840**(1): 29−36

[14] Wuhrer M, Koeleman C A, Hokke C H, *et al*. Protein glycosylation analyzed by normal-phase nano-liquid chromatography-mass spectrometry of glycopeptides. Anal Chem, 2005 , **77**(3): 886−894

[15] Kaji H, Shikanai T, Sasaki-Sawa A, *et al*. Large-scale identification of N-glycosylated proteins of mouse tissues and construction of a glycoprotein database, GlycoProtDB. J Proteome Res, 2012, **11**(9): 4553−4566

[16] McDonald C A, Yang J Y, Marathe V, *et al*. Combining results from lectin affinity chromatography and glycocapture approaches substantially improves the coverage of the glycoproteome. Mol Cell Proteomics, 2009, **8**(2): 287−301

[17] Song E, Zhu R, Hammoud Z T, *et al*. LC-MS/MS quantitation of esophagus disease blood serum glycoproteins by enrichment with hydrazide chemistry and lectin affinity chromatography. J Proteome Res, 2014, **13**(11): 4808−4820

[18] Makarov A, Denisov E, Kholomeev A, *et al*. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Anal Chem, 2006, **78**(7): 2113−2120

[19] Malerod H, Graham R L, Sweredoski M J, *et al*. Comprehensive profiling of N-linked glycosylation sites in HeLa cells using hydrazide enrichment. J Proteome Res, 2013, **12**(1): 248−259

[20] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol, 2008, **26** (12):1367−1372

[21] Hao P, Ren Y, Alpert A J, *et al*. Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. Mol Cell Proteomics, 2011, **10**(10): O111 009381

[22] Palmisano G, Melo-Braga M N, Engholm-Keller K, *et al*. Chemical deamidation: a common pitfall in large-scale N-linked glycoproteomic mass spectrometry-based analyses. J Proteome Res, 2011, **11**(3):1949−1957

[23] Lewandrowski U, Zahedi R P, Moebius J, *et al*. Enhanced N-glycosylation site analysis of sialoglycopeptides by strong cation exchange prefractionation applied to platelet plasma membranes. Mol Cell Proteomics, 2007, **6**(11): 1933−1941

[24] Cohen A M, Kostyleva R, Chisholm K A, *et al*. Iodination on tyrosine residues during oxidation with sodium periodate in solid phase extraction of N-linked glycopeptides. J Am Soc Mass Spectrom, 2012, **23**(1): 68−75

[25] http://weblogo.berkeley.edu/logo.cgi

[26] Wollscheid B, Bausch-Fluck D, Henderson C, *et al*. Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. Nat Biotechnol, 2009, **27**(4): 378−386

Comparison of Hydrazide Chemistry and Lectin Affinity Based Enrichment Methods for N-glycoproteomics*

XUE Peng[1,2], ZHOU Yue[1,2,3], DING Xiang[1], WANG Jun[1,2],
XIE Zhen-Sheng[1], YANG Fu-Quan[1]**

([1] *Institute of Biophysics, Chinese Academic of Sciences, Beijing* 100101, China;
[2] *University of Chinese Academic of Sciences, Beijing* 100049, China; [3] *Themofisher Scientific, Shanghai* 200120, China)

Abstracts Glycosylation is one of the most common and important post-translational modifications of proteins. Identification of large-scale N-linked glycoprotein is a very important aspect in glycoproteomics research. The N-glycopeptide enrichment is a key step in high-throughput identification of N-glycosylation site. Lectin enrichment and hydrazide chemistry are the two widely used N-glycopeptides enrichment methods. Each method can only enrich certain types of glycopeptides. It has been reported that the two methods are highly complementary, but few studies compared overlaps of glycosties from the two methods. In this paper, using HepG2 cells, we systematically compared the performance of hydrazide chemistry and lectins enrichment methods. The results showed that although the hydrazide method with glycopeptides enrichment efficiency of 76.7%, far higher than the 54.6% lectin enrichment method, 825 glycoprotein and 1 879 N-glycosylation sites identified with the lectin method was significantly more than 522 glycoprotein and 1 014 glycosylation sites enriched by the hydrazide method. Moreover, the two methods did not show significant complementary, together, only 853 glycoproteins and 1 959 N-glycosylation sites were identified. The overlapping results of identified N-glycosylation sites and N-glycoproteins from the two methods show that lectins enrichment method was better than hydrazide chemistry method.

Key words glycoproteomics, lectin affinity enrichment, hydrazide chemistry enrichment
DOI: 10.16476/j.pibb.2014.0324

---

*This work was supported by a grant from National Basic Research Program of China (2012CB966803, 2010CB833703) and the Scientific Instruments Development and Technical Innovation Program of Chinese Academy of Sciences(yg201001).

**Corresponding author.
Tel: 86-10-64888581, E-mail: fqyang@ibp.ac.cn

Received: January 20, 2015　Accepted: March 17, 2015